AP Statistics – 8.2	Name: **KEY**
Goal: Understanding Confidence Interval (CI) for Proportions	Date:

I. **<u>Definitions for</u>: Confidence Intervals for Proportions**

1) What are the 3 conditions that must be checked before finding a CI for proportions?

- **RANDOM: SRS or randomized sampling design**
- **INDEPENDENT: The population is at least 10 times as large as the sample.**
- **NORMAL: Number of successes $(n\hat{p})$ and failures $(n\hat{q})$ in the sample are both at least 10.**

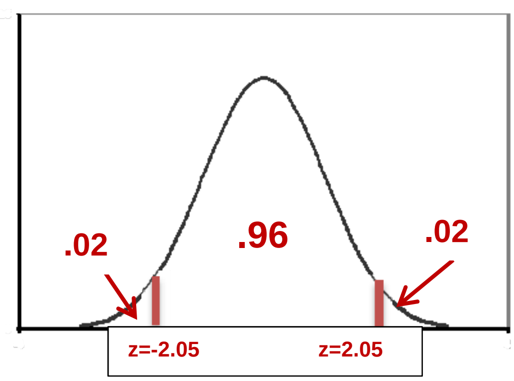
2) <u>Critical value:</u>

a) What statistic is used for the critical value to find a CI for proportions?

- For **normal distributions**, the critical value is **z***.

b) Find the critical value for a 96% confidence Interval.

Sketch the graph.



Use calculator to find the critical value. Explain.

- **invNorm(.02,0,1)= -2.0537**

  z* = ∓ 2.05

  (typically round 2 decimal)

b) Memorize the most common critical values. Find critical values for:

- 90% → 1.00-.90=.10/2=.05 → invNorm(.05,0,1) → $z^* = \mp 1.645$
- 95% → 1.00-.95=.05/2=.025 → invNorm(.025,0,1) → $z^* = \mp 1.96$
- 99% → 1.00-.99=.01/2=.005 → invNorm(.005,0,1) → $z^* = \mp 2.58$

3) What is the point estimator/estimate for proportions?

- Point Estimator is the statistic $\hat{p}$.
- The value of $\hat{p}$ is the Point Estimate.

4) What is the standard deviation/standard error for proportions?

Sampling distribution Point estimate

- standard deviation ↔ Standard error = $\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$
- **Standard error of the sample proportion $\hat{p}$ describes how far $\hat{p}$ will be from p, on averae, in repeated SRSs of size n.**

5) What is the margin of error for proportions?

- Margin of error $me = z^* \cdot \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$ **tells how close the estimate tends to be to the unknown parameter in repeated random sampling.**

6) What is formula to find CI for proportions? $\hat{p} \mp z^* \cdot \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$

II. Conditions for Confidence Intervals for Proportions (CYU on page 487):

1) Random NOT MET. This is a convenience sample.

   Independent A large high school has at least 10(100) = 1,000 students.

   Normal $n\hat{p} = 17$ and $n(1-\hat{p}) = 83$ are both at least 10.

2) Random MET. Random sample.

   Independent MET. The sample [10(25) = 250] is less than 10% of the population [1,000's of bags of chips produced].

   Normal NOT MET. $n\hat{p} = 25(3/25) = 3$ is not at least 10.

III. Create Confidence Intervals for Proportions (CYU on page 490):

1) Population: U.S. college students.

   Population parameter: true proportion who are classified as binge drinkers.

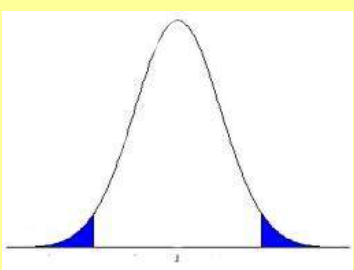
2) Random MET√ Random sample.

   Independent MET√ There are at least 10(10904)= 109,040 U.S. college students.

   Normal MET√ $n\hat{p} = 2486$ *and* $n\hat{q} = 8418$ is at least 10.

3) 99%→ 1.00-.99=.01/2=.005 → invNorm(.005,0,1) → $z^* = \mp 2.58$

   Show work highlighted in yellow!!!!



$\hat{p} = \frac{2486}{10904} = .228$

$.228 \mp 2.58 * \sqrt{\frac{.228(.772)}{10904}}$

$.228 \mp .010 \rightarrow (.218\ ,\ .238)$

Note:

$SE(\hat{p}) = \sqrt{\frac{.228(.772)}{10904}} = .004$

CHECK with Calc: [stat] [tests] [A: 1PropZInterval] x=2486 n=10904 C-Level=.99

Calc. Provides: (.21764, .23834) $\hat{p} = .228$

4) We are 99% confident that the interval from 0.218 to 0.238 captures the actual proportion of U.S. college students who are binge drinkers.

IV. Confidence Intervals for Proportions Examples:

Example #1 ***"Teens Say Sex Can Wait"*** The Gallup Youth Survey asked a random sample of 439 U.S. teens aged 13 to 17 whether they thought young people should wait to have sex until marriage. Of the sample, 246 said "Yes." Construct and interpret a 95% confidence interval for the proportion of all teens who would say "Yes" if asked this question.

a) Check conditions:

b) Find the z value:

c) Find $\hat{p}$:

d) Calculate the confidence interval:

e) Conclude (in context):

Key

EXAMPLE: Teens Say Sex Can Wait

Confidence interval for *p*

You must ALWAYS name the inference method by name or by formula
Name of this CI: 1-Sample Z-Interval for Proportions (p)

- **Check conditions:**
- **Random**: Gallup surveyed a random sample of 439 U.S. teens.
- **Normal**: We check the counts of success and failures

$$n\hat{p} = 439 * .56 = 246 \text{ and } n(1-\hat{p}) = 439 * .44 = 193$$

The counts of successes and failures are both ≥ 10

- **Independent**: Since Gallup sampled without replacement, we need to check the 10% condition. At least 10(439) = 4390 U.S. teens aged 13 to 17.

- **Find the 95% z value:** Calculator command: invNorm(.025,0,1) = -1.96 or invNorm(.975,0,1) = 1.96

- **Find**: $\hat{p}$ =246/439=.56

- **Calculate the confidence interval:** $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.56 \pm 1.96\sqrt{\frac{(0.56)(0.44)}{439}} = 0.56 \pm 0.046$

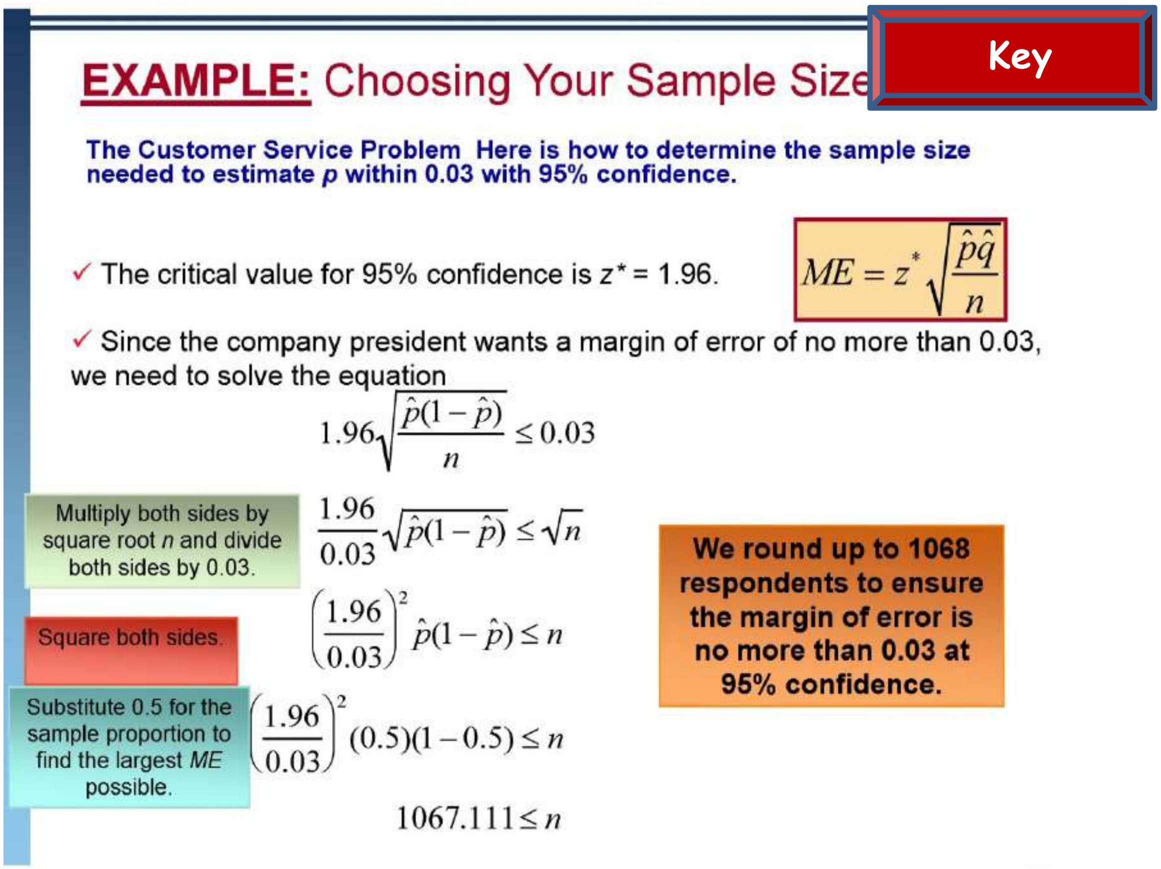
$= (0.514, 0.606)$

- **Conclude (in context)**: We are 95% confident that the interval from .514 to .606 captures the true proportion of 13-to 17-year-olds in the United States who would say that teens should wait until marriage to have sex.

V. Sample Sizes for Proportions Examples:

Example #2 A company has received complaints about its customer service. The managers intend to hire a consultant to carry out a survey of customers. Before contacting the consultant, the company president wants some idea of the sample size that they will be required to pay for. One critical question is the degree of satisfaction with the company's customer service, measured on a 5-point scale. The president wants to estimate the proportion p of customers who are satisfied (that is, who choose either "satisfied" or "very satisfied," the 2 highest levels on the 5-point scale).

The president wants the estimate to be within 3% (.03) at a 95% confidence level. How large a sample is needed?



CYU on page 494:

1) 95% CL: $1.96 * \sqrt{\frac{.8(.2)}{n}} \le .03$

$(1.96*.4)/.03 \le \sqrt{n}$

$682.95 \le n$ (n=683; always round up to ensure ME)

2) 99% CL: $2.58 * \sqrt{\frac{.8(.2)}{n}} \le .03$

$(2.58*.4)/.03 \le \sqrt{n}$

$1183.36 \le n$ (n=1,184; increasing the CL means sample size increases)